

## RESEARCH ARTICLE

# Comparative proteomic and transcriptional profiling of a bread wheat cultivar and its derived transgenic line overexpressing a low molecular weight glutenin subunit gene in the endosperm

Federico Scossa<sup>1, 2</sup>, Debbie Laudencia-Chingcuanco<sup>2</sup>, Olin D. Anderson<sup>2</sup>, William H. Vense<sup>2</sup>, Domenico Lafiandra<sup>1</sup>, Renato D'Ovidio<sup>1</sup> and Stefania Masci<sup>1</sup>

<sup>1</sup> Department of Agrobiological and Agrochemistry, University of Tuscia, Viterbo, Italy

<sup>2</sup> U.S. Department of Agriculture – ARS, Western Regional Research Center, Albany, CA, USA

We carried out a parallel transcriptional and proteomic comparison of seeds from a transformed bread wheat line that overexpresses a transgenic low molecular weight glutenin subunit gene relative to the corresponding nontransformed genotype. Proteomic analyses showed that, during seed development, several classes of endosperm proteins were differentially accumulated in the transformed endosperm. As a result of the strong increase in the amount of the transgenic protein, the endogenous glutenin subunit, all subclasses of gliadins, and metabolic as well as chloroform/methanol soluble proteins were diminished in the transgenic genotype. The differential accumulation detected by proteomic analyses, both in mature and developing seeds, was paralleled by the corresponding changes in transcript levels detected by microarray experiments. Our results suggest that the most evident effect of the strong overexpression of the transgenic glutenin gene consists in a global compensatory response involving a significant decrease in the amounts of polypeptides belonging to the prolamin superfamily. It is likely that such compensation is a consequence of the diversion of amino acid reserves and translation machinery to the synthesis of the transgenic glutenin subunit.

Received: September 6, 2007

Revised: February 14, 2008

Accepted: March 18, 2008

**Keywords:**

Genetically modified crops / Microarray / Seed proteome / Wheat

## 1 Introduction

Wheat is the world's most important crop and products based on this grain constitute the staple food in the majority of world countries. Due to its economic impact, and with the

aim of improving important traits, wheat has always been a target for conventional breeding programs. In recent years, wheat has been shown to be amenable to genetic transformation, and several efforts are being made to transform it with additional genes of agronomic importance related to end-use quality [1].

Bread and pasta are the most common wheat-based products, and their quality is mainly determined by the rheological properties of water-flour doughs, which correlate with the quantity, composition, and structural properties of the gluten proteins. Gluten is a highly complex mixture of polypeptides that amounts to about 80% of the total proteins in mature caryopses [2].

Gluten proteins are classically divided into gliadins and glutenins. Gliadins are monomeric proteins, and can be further classified as  $\alpha$ -,  $\beta$ - (now considered the same protein

**Correspondence:** Professor Stefania Masci, Department of Agrobiological and Agrochemistry, University of Tuscia, Via S. Camillo de Lellis snc, 01100 Viterbo, Italy

**E-mail:** masci@unitus.it

**Fax:** +39-0761-357238

**Abbreviations:** CM, chloroform/methanol; DE, differentially expressed; DPA, days postanthesis; FDR, false discovery rate; GM, genetically modified; HMW-GS, high molecular weight glutenin subunit; JA, jasmonic acid; LMW-GS, low molecular weight glutenin subunit

class, named  $\alpha$ -type),  $\gamma$ -, and  $\omega$ -gliadins based on their mobility in PAGE gels at acidic pH [3]. Gliadins form only intramolecular disulfide bonds, or, in the case of  $\omega$ -gliadins, because of the absence of cysteine residues, no disulfide bonds. The glutenins, on the other hand, consist of polymeric proteins stabilized by both intra- and intermolecular disulfide bonds. The glutenin polymers can reach up to several million Daltons, and polymer size, abundance and composition determine the end-use quality of wheat [4]. Upon reduction of disulfide bonds, single polypeptides are released from the glutenin polymers, and can be broadly classified into high molecular weight glutenin subunits (HMW-GS) and low molecular weight glutenin subunits (LMW-GS). The HMW-GS range from 70 000 to 90 000, whereas the LMW-GS range from 20 000 to 45 000. LMW-GS can be further divided on the basis of pI and molecular weight [5], into the B, C, and D groups. The majority of LMW-GS are present in the B group, have distinctive N-terminal sequences [6], and may be considered “typical” LMW-GS. The C group contains mainly polypeptides with  $\alpha$ -type or  $\gamma$ -gliadin type sequences, and their incorporation in the polymeric fraction of gluten is likely due to mutations altering the number of cysteine residues [7]. The D-group of LMW-GS consists of polypeptides with typical  $\omega$ -gliadins sequences that are similar to the C-type LMW-GS, in that their presence in the glutenin polymers is likely due to mutations resulting in the presence of at least a single cysteine residue [8–10].

In addition to the gluten proteins, wheat grains also contain nongluten proteins, which are involved in the biochemical processes that take place during grain development and germination. These nongluten proteins constitute only a minor fraction in terms of quantity (20% of the total endosperm proteins). The most abundant of these are evolutionarily related to the wheat prolamins and are part of the prolamins superfamily. The nongluten protein fraction includes water- and salt-soluble enzymes active in carbohydrate metabolism, protein synthesis, transport, and various stress responses; some of them, notably those soluble in salt solutions and in chloroform/methanol (CM-like proteins), are mainly polypeptides with a proposed role in plant defense against pathogens [11–13]. Recent advances have made possible the selective extraction and characterization of metabolic proteins using a combination of 2-DE and MS [14–16].

Previous studies reported the production of transgenic wheats overexpressing several HMW-GS genes in bread wheat [17, 18] as well as LMW-GS in durum wheat [19, 20]. In all cases, modifications of visco-elastic properties of the derived doughs were observed [21]. Because LMW-GS are also important determinants of the rheological properties of bread wheat doughs [22], bread wheat with additional copies of LMW-GS genes was generated [23]. Investigations currently in progress [24], based on the parallel fluorescence *in situ* hybridization (FISH) analyses of untransformed and genetically modified (GM) genotypes, along with a mono-

somic/nullisomic analysis, have shown that the LMW-GS transgenes, along with the BAR marker genes, are both integrated in a single locus, confirming the previously formulated hypothesis that the LMW-GS overexpression is due to a multiple copies insertion event [23].

With the aim of evaluating the consequences of transgene presence and activity on the expression of the other classes of endosperm genes, we employed a genomic approach to compare transcript and protein synthesis during the period of seed filling until maturation. Seed development, from pollination to mature desiccated grain, generally occurs over 45 days in bread wheat under standard environmental regimes. Proteins and starch, the main seed storage reserves, usually accumulate in a linear fashion between 10 and 35 days postanthesis (DPA) [25]. We chose to focus our study over this period, during which the transgenes are expected to be transcribed, and where most of the gene expression differences are likely to occur.

The data presented here provide a definition of genome and proteome responses to genetic transformation and overexpression of a transgenic LMW-GS gene in the developing seeds of a transgenic bread wheat line. They show that the main consequence is a decreased gene expression and synthesis of most of the endosperm proteins belonging to the prolamins superfamily. It seems likely that such mechanism of compensation is a consequence of the diversion of amino acids reserves and translation machinery to the synthesis of the transgenic glutenin subunit.

## 2 Materials and methods

### 2.1 Plant materials

The transformed *Triticum aestivum* L. (bread wheat) cultivar Bobwhite (carrying the 1B-1R translocation) over expressing an LMW-GS gene, along with its corresponding control genotype is described in ref. [23]. Briefly, the construct pLMWF23A used for biolistic transformation of immature embryos contains a full length LMW-GS coding sequence isolated from the bread wheat cultivar Cheyenne under the control of its own promoter [26]. We expect that the LMW-GS transgene, like the endogenous LMW-GS genes, is subject to developmental regulation and it is expressed exclusively in the starchy endosperm cells during the mid and the late stages of seed development [2]. On the basis of the molecular weight, pI and structure, the transgene-encoded protein belongs to the B-type of LMW-GS [5, 6, 23].

The transgenic genotype used corresponds to T<sub>7</sub> and T<sub>8</sub> generations, and thus is assumed to be completely homozygous. Seeds of the wild-type and the transformed genotypes seeds were germinated in Petri dishes at room temperature (in November 2003), transferred in Jiffy pots when the seedlings reached approximately 3 cm in height, and kept at 4°C for 4 wk. After this period, the plants were individually transplanted in 14 cm × 18 cm (diameter × height)

pots and grown in a greenhouse at ambient temperature until full maturity (May 2004). A null segregant line of the transgenic genotype (a line selected for the absence of transgene overexpression in the first segregating generations, as well as for the absence of the marker gene) was also grown in the greenhouse under identical conditions, and used for the proteomic comparisons reported here.

To stage developing seeds used for either the proteomic or microarray analyses, heads of plants grown in the same year were tagged when the first anther was exposed (0 DPA). Developing seeds were then harvested at 10, 20, 30, and 35 DPA by flash-freezing the whole spikes in liquid nitrogen and collecting the caryopses on dry ice. Frozen developing caryopses were stored at  $-80^{\circ}\text{C}$  until RNA or protein extraction. After careful observation of the caryopses morphology, we decided to split the 10 DPA time-point into two subgroups: “early” and “late,” according to the different length across the seed longitudinal crease, and thus to the volume of the endosperm.

For biological replicates, we pooled the material collected from spikes harvested from different plants. We obtained enough material to perform two (10 DPA early) and at least three (10 late, 20, 30 and 35 DPA) biological replicates, from which independent RNA isolations and protein extractions were performed.

## 2.2 Microarray profiling

A wheat cDNA array with 10 800 features representing 7835 genes [27] was used in the experiments. The arrays contain inserts from cDNA clones generated from 48 cDNA libraries generated by a consortium [28] funded by the US National Science Foundation (“The Structure and Function of the Expressed Portion of the Wheat Genome”). A majority of the probes on the array (5817) have been mapped on the hexaploid wheat genome [29].

A direct two-sample comparison experimental design, including dye swap hybridizations, was applied [30]. For each time-point, total RNA from both the wild-type and the transgenic genotype was reverse-transcribed, labeled, and the pooled target solution was then applied to the cDNA array. All hybridizations were technically replicated (involving the same RNA samples) by swapping of the dyes. The whole experiment comprises 28 cDNA array slides.

To assess the array reproducibility, control experiments in which the same RNA sample was labeled with Alexa 647 and Alexa 555 (Molecular Probes) were performed (“yellow test,” data not shown). Using the normalization and data analysis procedures briefly described below and with full details in the Supporting Information Appendix S1 (MIAME checklist), none of the genes were found to be statistically differentially expressed (DE) on the yellow test (with a false discovery rate (FDR)-corrected  $p$ -value less than 0.01). Globally, only 2.4% of the clones (260 clones out of 10 800) showed  $>1.5$ -fold variation; and only 0.64% (70 clones out of 10 800) showed  $>2$ -fold variation. These data are comparable

to those reported in ref. [31] regarding the technical degree of variation of a cDNA array.

Extraction and labeling of total RNA from developing caryopses of the wild-type and transgenic genotypes were performed as described [32]. Briefly, RNA was indirectly labeled with Alexa 555 or Alexa 647 fluorophores (Invitrogen, Carlsbad, CA) and hybridizations were carried out using the Pronto! Hybridization kit reagents (Corning) following the instructions supplied by the manufacturers. Array slides were scanned using a GenePix 4000B Scanner (Molecular Devices, Sunnyvale, CA), equipped with GenePix Pro 6.0 data acquisition software (Axon Instruments). Array data normalization and detection of DE genes were carried out using the software package LIMMA (Linear Models for Microarray Analysis) [33, 34] loaded into the programming language R v. 2.2.1 [35, 36]. Selection of the DE clones was based on the  $p$ -values associated with the moderated  $t$ -statistics ( $\alpha = 0.01$ ) [37], after FDR correction for multiple comparison [38]. To determine the number of DE tentative unique genes encoded by the DE clones, the ESTs sequences were assembled into tentative contigs using the software Phrap (<http://www.phrap.org/phredphrapconsed.html>) as previously described [27].

The data related to the wild-type and to the transgenic genotype were normalized and averaged to obtain, for each clone at each time-point, a mean-absolute value of the expression level. For each clone, these values were then divided by the expression value of the first time-point, and ratios were then rescaled by setting the lowest ratio to 1 and plotted. This data presentation procedure has been shown to be better in visualizing the magnitudes of the variation in the gene expression levels [31]. For the purposes of data presentation a clone was considered reliable if its average signal intensity was greater than 200 intensity units and more than twice the average background intensity, in at least one channel [39].

Full details of the experimental design are summarized in the Supporting Information Appendix S1 (MIAME checklist), which is compliant to the MIAME checklist format [40].

## 2.3 Proteomic profiling

### 2.3.1 Metabolic and CM-like proteins

Flours (150 mg) from the wild-type genotype and from the transgenic line, derived from pools of mature seeds, were used for selective extraction of CM-like and metabolic proteins according to the procedure reported in ref. [16]. At least three different biological replicates were performed, both for metabolic and CM-like proteins, consisting in separate extractions from different flour samples.

To detect quantitative variations in the level of individual proteins, equal volumes of extracts made from equal masses of flours were loaded on the 2-D gels, rather than the same amount of proteins [41]. This was decided on the basis of the significant difference in protein contents between the two

genotypes, higher in the transgenic line with respect to the untransformed genotype [23], that would not allow to detect any difference in case the same amount of protein is loaded.

Samples were focused according to the following steps: 12 h rehydration at 20°C, then 200 V for 30 min, 500 V for 30 min, 1000 V for 2 h, 5000 V for 2 h with a final step of 8000 V for 5 h. Intensity was limited to 50  $\mu$ A/strip. Equilibration of IPG strips was performed as previously described [14, 16]. Second dimension was carried out at 30 mA/gel, with cooling at 10°C, and stopped 5 min after the tracking dye was out of the gel. Gels were stained overnight with CBB G-250 according to [42], destained for 1 h in distilled water and images were acquired immediately with a calibrated scanner at 300 dot *per* inch with the same settings for all gels (16-bit grayscale pixel depth). ImageMaster 2D v. 3.1 software (Amersham Biosciences) was used to match and analyze the gel images.

Spot detection and background subtraction (method: average on boundary) were made by the software with the same settings for all gels, and the detected spots were manually verified.

The 2-D gels were analyzed in order to quantify the global abundance of metabolic and CM-like protein by subtracting background from the total raw spot quantities. The values obtained were submitted to a two-tails Student's *t*-test to evaluate the hypothesis of a significant difference existing in the global amount of metabolic and CM-like proteins between the wild-type and the transgenic genotype.

A second type analysis of 2-D gels was made to identify protein spots that were differentially accumulated in the metabolic protein preparations. In this case, raw images were treated as previously described for spot detection and background subtraction – landmarks (user's matches) were added to improve matching efficiency. Volumes of the detected spots were normalized (total spot volume normalization) by dividing by the sum of spot volumes for all detected spots and then multiplied by 100 to obtain spot percentage (normalized spot volume). Student's *t*-statistics and fold changes for detecting differentially synthesized proteins were calculated using the normalized spot volumes. Statistically significant differentially synthesized protein spots in metabolic protein preparations showing a threefold or greater variation in normalized spot volume between the transgenic and the wild-type genotype were selected for MS/MS analysis.

### 2.3.2 Total gluten proteins

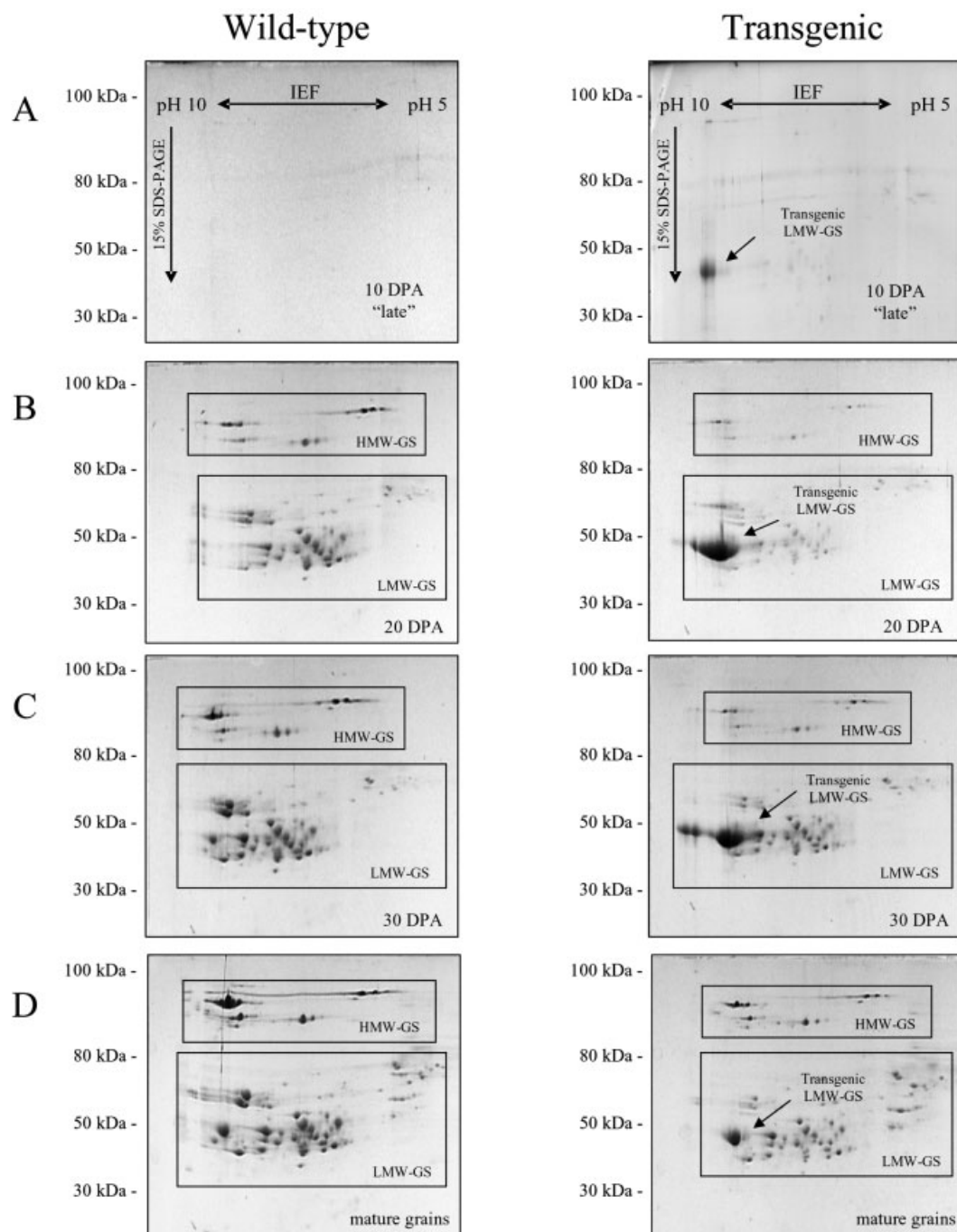
Total gluten proteins were extracted from pools of mature grains and from pools of developing caryopses according to the procedure reported in ref. [14] on the residual pellets obtained after the isolation of the metabolic proteins. Although this procedure extracts both gliadin and glutenin subunits, 2-DE allows to separate efficiently mainly HMW-GS and LMW-GS of B-type, because gliadins and C-type LMW-GS comigrate extensively. For this reason, these two latter classes of proteins have been analyzed separately.

Analysis by 2-DE was performed at least in biological triplicate. Pools of mature and developing seeds at 10 DPA late, 20 DPA, and 30 DPA were crushed to a fine powder in liquid nitrogen, immediately freeze-dried, and 200 mg were used for the extraction of gluten proteins. Pellets of gluten proteins obtained were dissolved and separated by means of 2-DE (IEF  $\times$  SDS-PAGE) according to the procedure reported above for metabolic and CM-like proteins. In this experiment, equal volumes of extracts made from equal masses of plant material were loaded on the 2-D gels [41]. Gel staining, image acquisition, background subtraction, and spot detection were performed as described above with the same settings for all gels.

In order to allow comparisons between the different gels, and to correct for subtle experimental variations, the volumes of the detected spots in the gels were normalized as reported above (total spot volume normalization). The normalized volumes of the spots detected in the HMW-GS and LMW-GS regions (boxed in Fig. 1) were used to calculate the ratio HMW-GS/LMW-GS and to plot the histogram bars reported in Fig. 2.

### 2.3.3 Gliadins

Pools of immature caryopses or mature seeds from the wild-type, transgenic and null segregant line were separately crushed and gliadins were extracted from 10 mg of flour with a buffer containing 10% v/v dimethylformamide and 1  $\times$  Protease Inhibitor Cocktail (Sigma), in a 1–5 (mg/ $\mu$ L) ratio, for 2 h at room temperature. After 10 min centrifugation at 13 000 rpm, aliquots of the same volume of the supernatants were dried down in a SpeedVac or loaded on a Acid-PAGE gel for immediate analysis using aluminum lactate buffer at pH 3.2, according to ref. [43], with minor modifications. To analyze the pattern of gliadins extracted from mature grains in more detail, 2-DE (Acid-PAGE  $\times$  SDS-PAGE) was performed on gliadin fractions extracted from at least three biological replicates of the wild-type, transgenic, and null segregant line. In this case, equal volumes of gliadin extracts made from equal masses of seeds were loaded on the 2-D gels. The procedure used is as reported in ref. [7], but mini-gels (BioRad, Hercules, CA) were used both in the first and the second dimension. Differences between the present and the published procedure thus include first dimension prerun conditions that was performed at 16 mA/gel for 50 min, whereas run was carried out at 12 mA/gel and stopped 5 min after the dye reached the bottom of the gel. Gels were stained overnight according to ref. [42]. The second-dimension electrophoresis was performed in the same apparatus used for the first dimension, and gels were the same size. The run was carried out at a constant voltage (200 V), and was stopped when the CBB used to stain the first dimension reached the bottom of the gel (about 50 min). Second dimension gels were stained overnight with colloidal Coomassie [42] and destained in distilled water. Gel images were acquired with a calibrated scanner at 300 dot *per* inches (16 bit pixel depth) with the same settings for all gels.

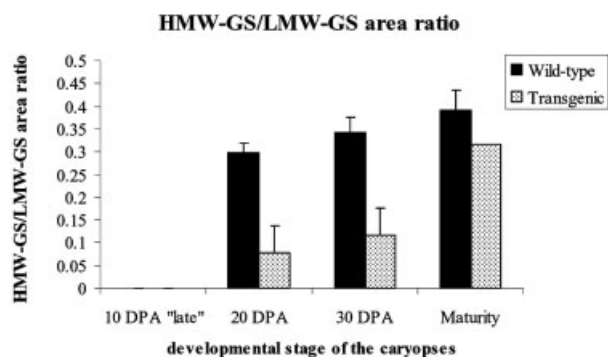


**Figure 1.** 2-DE (IEF  $\times$  SDS-PAGE) gels of total gluten protein preparations from the wild-type (left) and from the transgenic genotype (right) extracted at different time-points during seed development: (A) 10 DPA late; (B) 20 DPA; (C) 30 DPA; and (D) mature grains. Equal volumes of seed extracts derived from equal masses of seeds were loaded on the gels. The transgenic LMW-GS is indicated by the arrow. Regions including the HMW-GS and the LMW-GS are boxed.

Background subtraction (method: average on boundary), spot detection, matching, and normalization (method: total spot volume) were performed by the software ImageMaster 2D v. 3.1 (Amersham Biosciences) as described. Student's *t*-statistics and fold changes for detecting differentially syn-

thesized protein spots were calculated on the normalized spot volumes.

Electroblotting of the 2-D gels of gliadins was performed using an X Cell II blot module (Invitrogen). Sequi-blot PVDF membranes (BioRad) were wetted in methanol and rinsed



**Figure 2.** Histogram showing the comparison between the wild-type and transgenic HMW-GS/LMW-GS area ratios during seed development. At 10 DPA late gluten proteins are not yet detectable in the wild-type seeds, while the only subunit visible from the transgenic seeds is the transgenic LMW-GS. Error bars denote the SD. Where error bars are not visible, they are too small to be represented within the histogram scale. Calculation of the HMW-GS/LMW-GS area ratio is based on the normalized volumes of the detected spots included in the boxed regions in Fig. 1.

with deionized water for 5 min before soaking in electroblot buffer (10 mM 3-cyclohexylamino-1-propanesulfonic acid (CAPS), 10% v/v methanol, pH 11). Filter paper (Whatman 3 mm) was also soaked in electroblot buffer before use. After 2-DE, unstained gels were soaked in electroblot buffer for 5–10 min. Transfer was performed for 1 h at 4°C, at a constant voltage of 100 V. The membrane was rinsed with distilled water for 5 min before staining with 0.025% w/v CBB R-250 in 40% v/v methanol for 5 min. The membranes were destained for 5 min in 50% v/v methanol, briefly rinsed in distilled water, and allowed to air dry at room temperature. Statistically significant differentially accumulated protein spots showing a threefold or greater variation in normalized spot volume between the transgenic and the wild-type genotype were excised using a clean razor blade. N-terminal Edman amino acid sequencing was performed on the selected spots using an Applied Biosystems 492 Precise sequencer (Applied Biosystems, Foster City, CA) equipped with an online 140C Micro gradient HPLC system for identification of PTH-amino acids. The sample-containing PVDF membrane spots were coated with Biobrene (Applied Biosystems), dried and placed in to the reaction cartridge. The analysis cycle “PL PVDF Protein” that was supplied with the instrument was used for all samples. Sequence determination was carried out by visual inspection of the chromatograms and by examination of the data as analyzed by the Model 610A data analysis software.

### 2.3.4 C-type LMW-GS

In order to obtain fractions enriched in C-type LMW-GS, equal amounts of flour from at least three different biological replicates of mature grains of the wild-type, transgenic, and

null segregant line were used. We followed the procedure described by [7], consisting of different precipitation steps involving increasing concentrations of 1-propanol. C-type LMW-GS are better resolved by Acid-PAGE × SDS-PAGE rather than IEF × SDS-PAGE [7], because they do not present considerable differences in pIs. Acid-PAGE × SDS-PAGE was performed on the C-type LMW-GS preparations according to the procedure reported above for gliadins, with the addition of 0.5% w/v DTT to the solubilization buffer prior to sample loading on the first dimension gels. We followed the same procedures reported above for 2-D gels of gliadins regarding image acquisition, spot detection, normalization, and identification of differentially synthesized protein spots. Electroblotting and amino acid sequencing were also performed as described above.

### 2.3.5 MS

Selected protein spots from 2-DE gels were reduced, alkylated, digested with trypsin and identified using LC MS/MS as previously described [16].

#### 2.3.5.1 Protein identification

Tandem mass spectra were extracted from the QSTAR AnalystQS wiff files using a locally installed copy of MASCOT Daemon (<http://www.matrixscience.com/>) and the MS/MS samples were analyzed using MASCOT (Matrix Science, London, UK; version 2.1.04). MASCOT was set up to search a local version of a green plant database (containing 188 169 entries) assuming trypsin as the cleavage enzyme. The database was created by parsing the NCBI nonredundant (nr) database (03/10/2006) into taxonomic groupings using the Sequence Database Management Wizard from Genomic Solutions (Ann Arbor, Michigan). The local version of the green plant database contained nr-Arabidopsis-thaliana.fasta, nr-Other-Viridiplantae.fasta, and nr-Oryza-sativa.fasta and had appended to it protein sequences contained in the common Repository of Adventitious Proteins (04/06/2006) available from the Global Proteome Machine Organization at <ftp://ftp.thegpm.org/fasta/cRAP>. Using the postanalysis software package Scaffold (Proteome Software, Portland, OR) both MASCOT and X! Tandem [44] results were searched with a fragment ion mass tolerance of 1.00 Da and a parent ion tolerance of 1.00 Da. Iodoacetamide derivative of cysteine was specified in MASCOT and X! Tandem ([www.thegpm.org](http://www.thegpm.org); version 2006.04.01.2) as a fixed modification. Oxidation of methionine was specified in MASCOT as a variable modification. Oxidation of methionine and iodoacetamide derivative of cysteine were specified in X! Tandem as variable modifications.

#### 2.3.5.2 Criteria for protein identification

Scaffold (version Scaffold-01.05) was used to validate the combined MASCOT and X! Tandem MS/MS based peptide

and protein identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability as specified by the Peptide Prophet algorithm [45]. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least two identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm [46]. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

### 3 Results and discussion

In order to facilitate the comprehension of the data reported here, we will first present a summary of the results, both of the microarray and the proteomic analyses (Section 3.1). We will then present the details in Section 3.2, with data regarding storage protein and genes in Section 3.2.1, and metabolic proteins in Section 3.2.2. Because storage proteins (gluten) include a wide range of protein types (and relative coding genes), Section 3.2.1 has been further subdivided into three subsections, the first reporting data mostly on glutenins (Section 3.2.1.1), the second on a particular subclass of glutenin subunits, C-type LMW-GS (Section 3.2.1.2), and the third (Section 3.2.1.3) on gliadins.

#### 3.1 Effect of LMW-GS transgene overexpression on gene regulation during seed development

Our microarray analysis identified a total of 250 DE clones (from now on designated as “DE”) for all the five time-points analyzed, which covers the grain filling stage in seed development. To improve the level of confidence in identifying the clones that were up- or down-regulated, all clones with a FDR-corrected *p*-value of 0.01 or less (thus regarded as highly significant) were filtered again based on their expression fold-change ( $\geq \pm 2$ ). The clones passing the expression threshold were assembled into tentative unique DE genes

using the software Phrap as described previously [27, 32]. As shown in Table 1, no DE genes were detected at 10 DPA early, the first time-point analyzed. At this stage, the LMW-GS transcripts were not detectable. The onset of up-regulation of the LMW-GS transgenes is evident only starting from 10 DPA late and persisted until seed maturation (see below). The fact that seeds collected at the same time-point show different morphologies and different results as well, points out the need to take into consideration the grain developmental stage rather than the time-point itself.

The most evident consequence induced by genetic transformation with an LMW-GS transgene is a global down-regulation of gene expression during seed development. As shown in Table 1, the proportion of down-regulated unigenes in the transgenic genotype increased during seed development starting at 20 DPA, and reaching 100% at 35 DPA. This trend in down-regulation of genes in the transgenic genotype is paralleled, until 30 DPA, by the overexpression of several clones encoding the LMW-GS, very likely only the transgenic one, because of crosshybridization. Such crosshybridization is very likely due to the fact that the transgenic LMW-GS transcripts are expected to anneal to all LMW-GS probes, because of their high degree of sequence identity (>90%).

It is likely that this high overexpression of the transgenic LMW-GS gene requires an appropriate amount of transcription factors. Clones corresponding to transcription factors actually appear to be also overexpressed, in particular at 10 DPA late, when transcription of the LMW-GS transgene is maximum (see Supporting Information Appendix S2). Prolamin genes show similar patterns of temporal expression during seed development and they are thus expected to have common regulatory sequences [2, 47]. Hierarchical clustering (cluster of the genes *vs.* timepoint) calculated in the wild-type genotype showed extremely high correlation coefficients ( $R > 0.95$ ) between the endogenous glutenin genes (HMW-GS and LMW-GS) and some genes putatively encoding transcription factors. The genes encoding the same transcription factors were also found to be highly correlated, in the transgenic genotype, with the expression of the transgenic LMW-

**Table 1.** The DE unigenes detected by microarray analysis<sup>a)</sup>

Time-point	Replicated hybridizations <sup>b)</sup>	Number of DE unigenes	Up-regulated unigenes (transgenic/wild-type)	Down-regulated unigenes (transgenic/wild-type)	Proportion of down-regulated unigenes (transgenic/wild-type) <sup>c)</sup>
10 DPA early	4 (2)	0			
10 DPA late	6 (3)	31	25	6	19.3%
20 DPA	6 (3)	50	5	45	90%
30 DPA	6 (3)	78	4	74	94.9%
35 DPA	6 (3)	8	0	8	100%
Total	28 slides	167			

DE, differentially expressed.

a) Total number and proportion of statistically significant DE unigenes.

b) Number of biological replicates is indicated in parentheses.

c) Percentage of down-regulated unigenes (transgenic/wild-type) on the total number of DE unigenes.

GS transcript (see Supporting Information Appendix S3 “Transcript Profiles” for details). In fact, in the transgenic genotype, as the LMW-GS transcript level decreased (after its maximum peak at 10 DPA late), the transcript level of the transcription factors decreased as well. It thus seems reasonable that the accumulation of the transgenic LMW-GS transcript led to a significant depletion of transcription factors, causing the subsequent down-regulation of the other transcripts encoding the various prolamins.

The tentative contigs and singletons obtained from the assembly of the DE clones were all queried to the non-redundant database of NCBI using the algorithms BlastN and BlastX [48]. The putative identity was extracted and used to annotate each unigene. All DE unigenes identified, both those up- and down-regulated in the transgenic with respect to the wild-type genotype, were clustered in several functional categories and are summarized in Tables 2 and 3, respectively.

As expected on the basis of the very high sequence identity among LMW-GS genes [6], a high overexpression of several different LMW-GS clones has been detected by micro-

arrays in the transgenic genotype. This result is most likely due to crosshybridization between the LMW-GS array DNAs and the LMW-GS transgenic transcripts.

Among the down-regulated clones, the top-represented categories (Table 3) include storage and defense proteins. The clones encoding seed storage proteins (mainly HMW-GS and gliadins) were all down-regulated in the transgenic genotype, from the middle (*i.e.*, 20 DPA) to the late stages of seed development (*i.e.*, 35 DPA) (Figs. 3B–E). It should be noted that the endogenous LMW-GS were also down-regulated, as will be confirmed later in the proteomic analysis section (Fig. 1).

A high number of clones, including various classes of protease inhibitors, CM-like proteins, peroxidases, and lipid transfer proteins, which have all been attributed to plant defense [49], were down-regulated in the transgenic genotype.

Another class that appears down regulated in the transgenic overexpressing genotype includes genes involved in glutamine biosynthesis. Although these genes do not appear DE at the earlier stages (10 DPA), they show down regulation

**Table 2.** Functional classification of up-regulated unigenes detected by microarray analysis

Functional category	Representative unigene names	Maximum fold change <sup>a)</sup>	Time-point <sup>b)</sup>
Seed storage proteins	LMW glutenin subunit (transgene)	10	10, 20, and 30 DPA
Transport	Sarcoplasmic reticulum protein, SEC1 gene	7	10 DPA
JA biosynthesis	Oxophytodienoate reductase	7	10 DPA
Transcription and Translation	Zinc finger protein, putative transcription factor, putative chromosome decondensation factor, bZIP transcription factor, ribosomal proteins	13	10 and 20 DPA
Transcription	Zinc finger protein, bZIP transcription factor	3	30 DPA
Unknown	Several clones	5	10, 20, and 30 DPA

JA, jasmonic acid.

a) Fold change value refers to the maximum intensity ratio calculated as transgenic/wild-type detected over the five time-points analyzed.

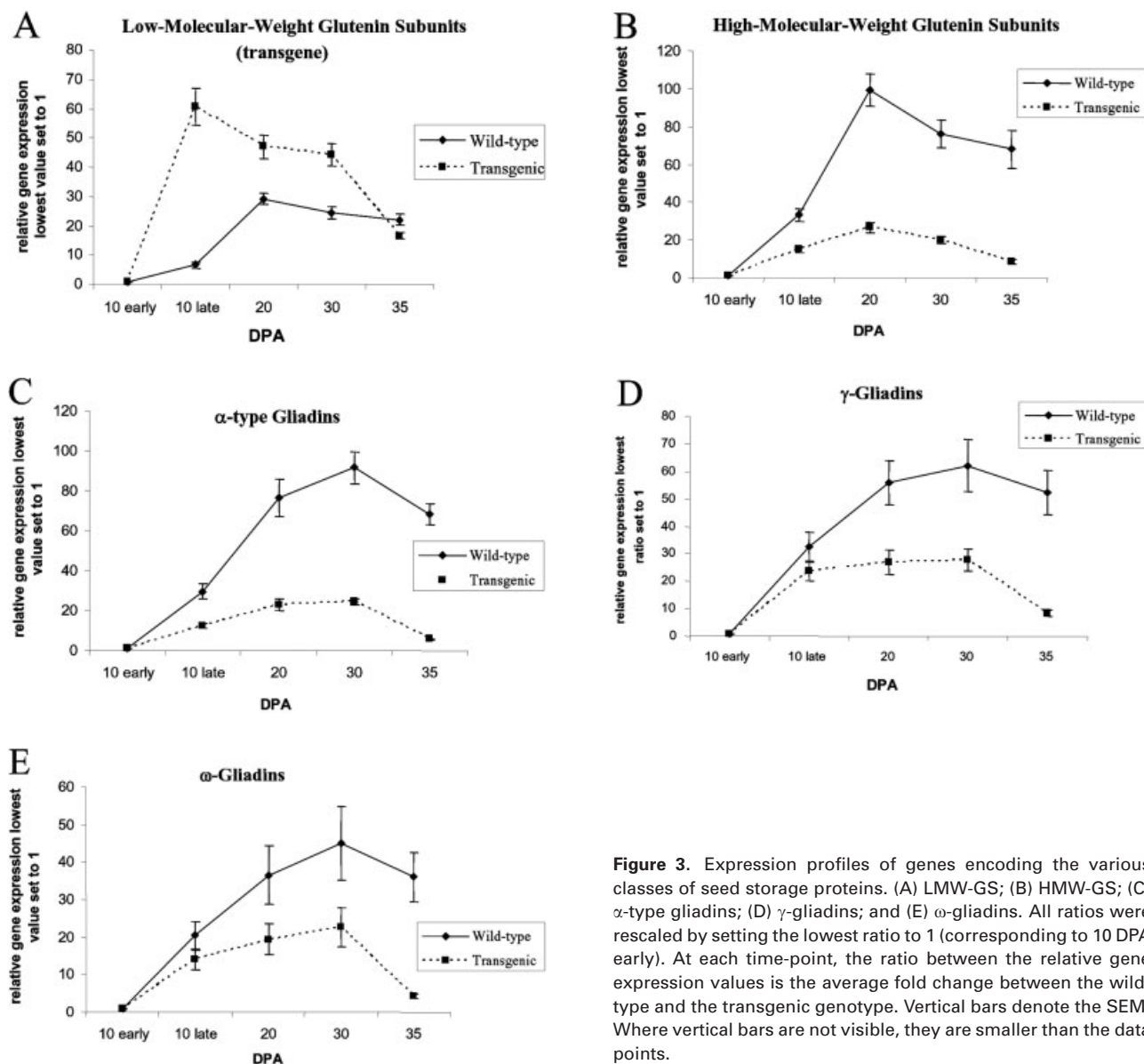
b) The column indicates when differential expression was detected.

**Table 3.** Functional classification of down-regulated unigenes detected by microarray analysis

Functional category	Representative unigene names	Maximum fold change <sup>a)</sup>	Time-point <sup>b)</sup>
Seed storage proteins	HMW glutenin subunit	–7	20, 30, and 35 DPA
	$\alpha$ -Type gliadins	–10	
	$\gamma$ -Gliadins	–7	
	$\omega$ -Gliadins	–8	
Defense	$\alpha/\beta$ -Amylase/trypsin inhibitors, CM soluble proteins (CM-like proteins), other protease inhibitors, peroxidases, lipid transfer proteins	–5	20, 30, and 35 DPA
Glutamine biosynthesis	Glutamine synthetase, glutaminyl-tRNA synthetase	–3	20, 30, and 35 DPA
Unknown	Several clones	–10	20, 30, and 35 DPA

a) Fold change value refers to the maximum intensity ratio calculated as wild-type/transgenic with inverted sign.

b) The column indicates when differential expression was detected.



**Figure 3.** Expression profiles of genes encoding the various classes of seed storage proteins. (A) LMW-GS; (B) HMW-GS; (C)  $\alpha$ -type gliadins; (D)  $\gamma$ -gliadins; and (E)  $\omega$ -gliadins. All ratios were rescaled by setting the lowest ratio to 1 (corresponding to 10 DPA early). At each time-point, the ratio between the relative gene expression values is the average fold change between the wild-type and the transgenic genotype. Vertical bars denote the SEM. Where vertical bars are not visible, they are smaller than the data points.

at the later stages, from 20 to 35 DPA. We hypothesize that this might be a consequence of an increased demand for the enzymes encoded by these genes at the earlier stages of development, when the transgenic LMW-GS, that is glutamine rich, starts to accumulate significantly (Fig. 3A). In the transgenic genotype, expression of these genes then decreased, likely for energy or metabolic reasons, thus leading to a down regulation of the other storage proteins, which are also glutamine rich. This aspect will be described in more detail in the following paragraphs.

Many genes putatively involved in transcriptional/translational processes and in the trafficking routes appeared strongly up-regulated. For some of them, especially those related to transcription, their differential expression per-

sisted until the later stages of development (Table 2). Microarray data indicate that the clone encoding the 12-oxophyto-dienoate reductase (OPR1, BE591113), an enzyme involved in the jasmonic acid (JA) biosynthesis, was strongly up-regulated in the transgenic genotype. Its transcript profile resembled, both in shape and magnitude of differential expression, the transcript profile of the transgenic LMW-GS (correlation coefficient  $[R] = 0.99$ ) (see Supporting Information Appendix S3 Transcript Profiles for details). Jasmonates are plant regulators able to induce the synthesis of storage and defense-related proteins in response to different biotic and abiotic stresses [50]. The observation that most of defense related genes appear down-regulated in the transgenic overexpressing genotype, suggests that the up-regula-

tion of the above-mentioned 12-oxophytodienoate reductase gene might be related to the increased LMW-GS synthesis, instead of solicited as a defense response.

Several clones showing significant changes have no known function. Bioinformatic analyses at the time of writing this manuscript did not provide any reliable annotation for these clones (lack of close similarity matches, gene ontology terms, and/or mapping information). This is consistent with the current status of genome annotation in wheat, and is of course a major issue in the complete biological interpretation of the gene expression datasets.

### 3.2 Gene expression profiles and protein accumulation of selected classes of proteins during seed development

As described in the Supporting Information Appendix S1 (MIAME checklist), the microarray dataset has been treated in order to perform a comparison between the different time-points, thus allowing the generation of graphs that give a more integrated vision of how a specific gene expression level changes across time. Many of the significant DE clones identified belong to gene families whose members share a high degree of identity at the level of nucleotide sequence. This is the case for seed-storage protein related genes (gliadins and glutenin subunits). To provide a unique transcript profile for each gene family, the gene expression profiles reported here are derived from pools of the mean intensity values from different array clones belonging to the same gene family.

The gene expression profiles of the main seed proteins were also compared at the proteomic level.

#### 3.2.1 Storage proteins

In the microarray analysis, it can be seen that genes corresponding to all the storage proteins (Fig. 3) in the wild-type genotype (HMW-GS, LMW-GS,  $\alpha$ -type,  $\gamma$ -, and  $\omega$ -gliadins) follow a similar pattern of expression, with glutenins, both HMW-GS and LMW-GS showing the peak of maximum abundance of transcripts earlier (at 20 DPA) than gliadins (at 30 DPA) (Fig. 3). This pattern of expression is in agreement with previous findings both at mRNA [51] and protein levels [52].

Comparison of the transgenic and the wild-type genotypes shows that there is no difference in the storage protein gene expression between the two genotypes at 10 DPA early; *i.e.*, the mRNA of all storage protein gene classes is not detectable. Significant differences were detected at the later stages in development (see details in the following Sections).

In the proteomic comparison, we first analyzed the total gluten proteins, which typically include both the HMW-GS and LMW-GS, along with the gliadins. However, the LMW-GS and gliadins partially comigrate on 2-D gels, and so we further investigated these two classes of proteins by using specific purification protocols.

Since the proteomic comparison involve different fractionation steps of the total seed proteins, the issue of a possible difference in the solubility properties of some of the protein classes between the two genotypes might arise. We have not found any evidence of changes of solubility of the different endosperm protein classes, also because the differences observed in gliadin and metabolic polypeptides amounts could be attributed to a different regulation of gene expression, as indicated by microarray analysis, rather than to solubility changes. Moreover, we had already observed [23] that the transgenic polypeptide forms mainly homopolymers, but it likely links also to endogenous glutenin subunits by means of the typical types of intermolecular disulfide bonds that form the gluten polymer.

#### 3.2.1.1 Glutenin proteins and genes

Starting with the LMW-GS protein class, encoded by the transgene, the gene expression profiles in the transgenic and in the untransformed genotypes (Fig. 3A) show a huge difference at 10 DPA late. The LMW-GS mRNAs of the transgenic genotype were ten-times more abundant than in the wild-type one, and were also significantly higher (about double the amount) at 20 and 30 DPA. The peak of maximum abundance of LMW-GS mRNAs is different between the two genotypes: in the transgenic line the LMW-GS transcripts had a maximum at 10 DPA late, whereas the wild-type reaches maximum at 20 DPA (Fig. 3A). This behavior might have different origins. It is possible that the construct containing the transgene does not include all the necessary control elements for the proper transcription, or the insertion site may affect expression. Both possibilities include the potential influence of higher order chromatin domain elements, such as MARs [53]. However, an alternative and simpler explanation of this shift in the peak of maximum transcript abundance, as well as of the apparently anticipated transcription and accumulation, is that the transgene might be one of the LMW-GS genes that are expressed earlier than others, because there are several genes coding for LMW-GS that are probably expressed at different times during the grain filling stage in the untransformed line. The presence of numerous transgene copies in the genome might thus account for this observation. We have in fact verified that multiple transgene copies are inserted in a single locus on the short arm of chromosome 5D [24]. On the basis of previously performed Southern blots [23] and FISH analyses [24], the estimated number of transgene copies is reasonably around one hundred, thus accounting for the observed differences both at transcriptional and translational levels, assuming that all the copies are active.

Proteomic analysis showed that the temporal modification of LMW-GS gene expression led to the accumulation of the transgenic subunit. Expression analysis by means of 2-D gels of gluten proteins extracted from seeds collected at 10 DPA late showed that the only detectable protein in the transgenic genotype is the transgenic LMW-GS itself,

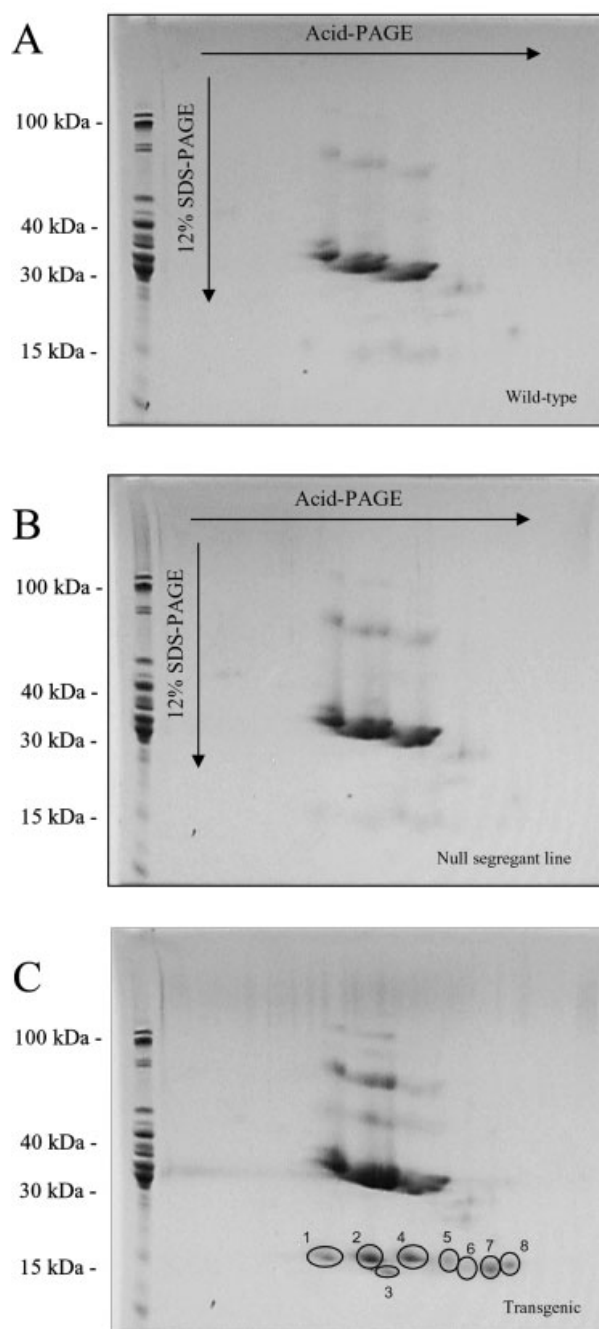
whereas no polypeptide has been detected in the wild-type cultivar (Fig. 1A). The prevalence of the transgenic LMW-GS (indicated by an arrow in Fig. 1) is particularly marked at 20 DPA (Fig. 1B), and still detectable at 30 DPA (Fig. 1C) and in the mature seed (Fig. 1D). It is important to stress that 2-D gels of gluten proteins were loaded with the same volume of protein extracts derived from equal masses of seeds, in order to reveal possible changes in the level of accumulation of individual polypeptides.

Regarding HMW-GS, their clones were also differentially regulated (Fig. 3B; Table 3). In the wild-type, HMW-GS transcripts have a similar pattern of expression to the LMW-GS genes: they show a drastic increase between 10 DPA late and 20 DPA, with a peak of maximum abundance at 20 DPA, and then their abundance decreases, when the seeds begin to desiccate. This trend is present also in the transgenic genotype, but the relative amounts are significantly different. Starting at 20 DPA, in the transgenic genotype, the transcripts for HMW-GS are in fact constantly down-regulated which accounts for the lower amount of HMW-GS polypeptides detected at 20, 30 DPA and in the mature seeds (Figs. 1B–D). The 2-D patterns reported in Fig. 1 allowed a parallel comparison between protein levels and gene expression profiles of glutenin subunits. Areas containing HMW-GS and LMW-GS spots are boxed in Fig. 1.

The dynamics of accumulation of the gluten proteins can be summarized looking at changes in the ratio between abundances of the HMW-GS and LMW-GS during seed filling (Fig. 2). This parameter had already been used to characterize the timing of seed storage proteins accumulation during development of the wheat caryopses [52, 54]. The HMW-GS/LMW-GS ratio was calculated taking into account all the normalized spot volumes of the HMW-GS and LMW-GS deriving from the expression analysis of the 2-D gels (boxed regions in Fig. 1). Figure 2 reports the histogram relative to this ratio. The HMW-GS/LMW-GS ratio increases in developing seeds in both genotypes analyzed. This observation is in agreement with the results reported earlier [52]. However, there is a strong difference in the absolute value of the HMW-GS/LMW-GS ratio. The ratio is constantly significantly lower in the transgenic genotype from 20 DPA until maturity (*t*-test *p*-values <0.05). This difference is due to the preferential accumulation of the LMW transgenic subunit at the earlier stages and to its high amount at the later stages of development (Figs. 1A and B).

### 3.2.1.2 LMW-GS of C type

The 2-D gels of C-type LMW-GS from the genotypes analyzed are shown in Fig. 4. Expression analysis of the 2-D gels identified as significantly differentially synthesized a group of fast moving polypeptides (around 15 kDa) present only in the transgenic genotype. These spots, circled in Fig. 4C, were submitted to N-terminal amino acid sequencing (Table 4),



**Figure 4.** 2-DE (Acid-PAGE × SDS-PAGE) of C-type LMW-GS from the wild-type (A), null segregant line (B), and from the transgenic genotype (C). A sample of total seed protein extract from the wild-type genotype was included as a reference (on the left of each image). Circled spots are exclusively expressed by the transgenic genotype and have been submitted to N-terminal sequencing (see also Table 4).

because no reliable data were obtained using MS, likely due to the known problems in obtaining suitable fragmentation patterns from these prolin-rich proteins [55].

**Table 4.** N-terminal sequences of DE spots from Fig. 4<sup>a)</sup>

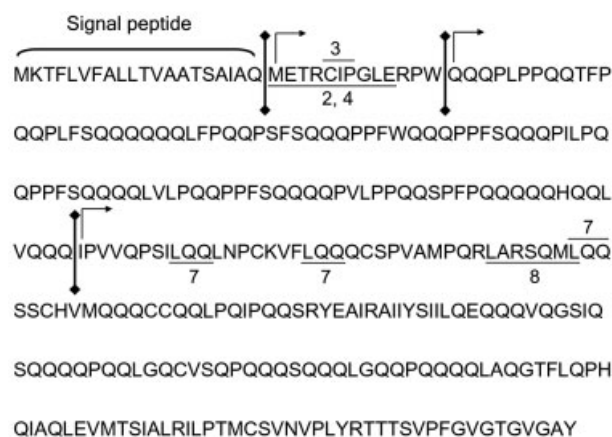
Spot no.	t-test <sup>b)</sup>	N-terminal sequence
1	<0.001	No sequence obtained
2	0.007	METRQIPGLE-
3	0.042	XXXXCIP-
4	0.005	METRQIPGLE-
5	0.011	XPTS(Q/P)(Q/P)(Q/G)L-
6	0.010	XQRSQMLVQ-
7	0.014	LQQ-
8	0.002	LARSQML-

a) N-terminal amino acid sequences obtained from the DE circled spots highlighted in Fig. 4C.

b) *p*-value to assess the differential expression. *t*-statistics were calculated on normalized spot volumes.

Spots numbers 2 and 4 corresponded to N-terminal peptides of typical LMW-GS. It is necessary to specify that, since the samples were not pyridylethylated, the glutamine in position 5 in peptides 2 and 4, might actually be a cysteine residue, as is in the transgenic LMW-GS (Fig. 5). Spot numbers 3, 7, and 8 correspond to internal peptides of the same protein type. If a cysteine residue is present in position 5, as it seems likely, it appears reasonable that fragments 2, 3, 4, 7, and 8 belong to the transgenic LMW-GS and might correspond to truncated transgenes that are generated as a consequence of the transformation process. Fragments of unexpected size putatively corresponding to partial sequences of the transgene have also been reported in several wheat lines transformed, by means of particle bombardment, with additional HMW-GS genes [56–61]. These fragments may also originate as a consequence of a proteolytic cleavage of the transgenic glutenin. Degradation of various transgenic products has been reported in reproductive and vegetative tissues of *Nicotiana tabacum*, *Arabidopsis thaliana*, corn, soy, alfalfa, and potato [62]. The observed alteration of the transgenic LMW-GS accumulation could lead to an imperfect assembly and misfolding of the protein in the ER lumen, therefore increasing the susceptibility of the transgenic glutenin to the proteases active in the secretory pathway [63, 64].

One of the sequences obtained for spot 5 (PTSQIQQI-), has been identified as an internal sequence of a protein similar to those of the zinc finger family (NP\_498846, *E*-value = 227). Several clones encoding zinc finger transcription factors (in particular BE423328 and BQ804784) showed to be strongly up-regulated in the transgenic genotype from 10 DPA late to 30 DPA (data not shown). The sequence obtained for spot 6 (QRSQMLV-) matches an internal predicted sequence of a hydrolase (ZP\_01111041, *E*-value = 126). However, spots 5 and 6 may also be products of rearranged LMW-GS transgenes. Thus, the novel DE spots identified in the transgenic C-type LMW-GS preparations may correspond to by-products of the overexpressed LMW-GS or derive from transgene rearrangements, rather than to *de novo* polypeptide synthesis (Fig. 5).

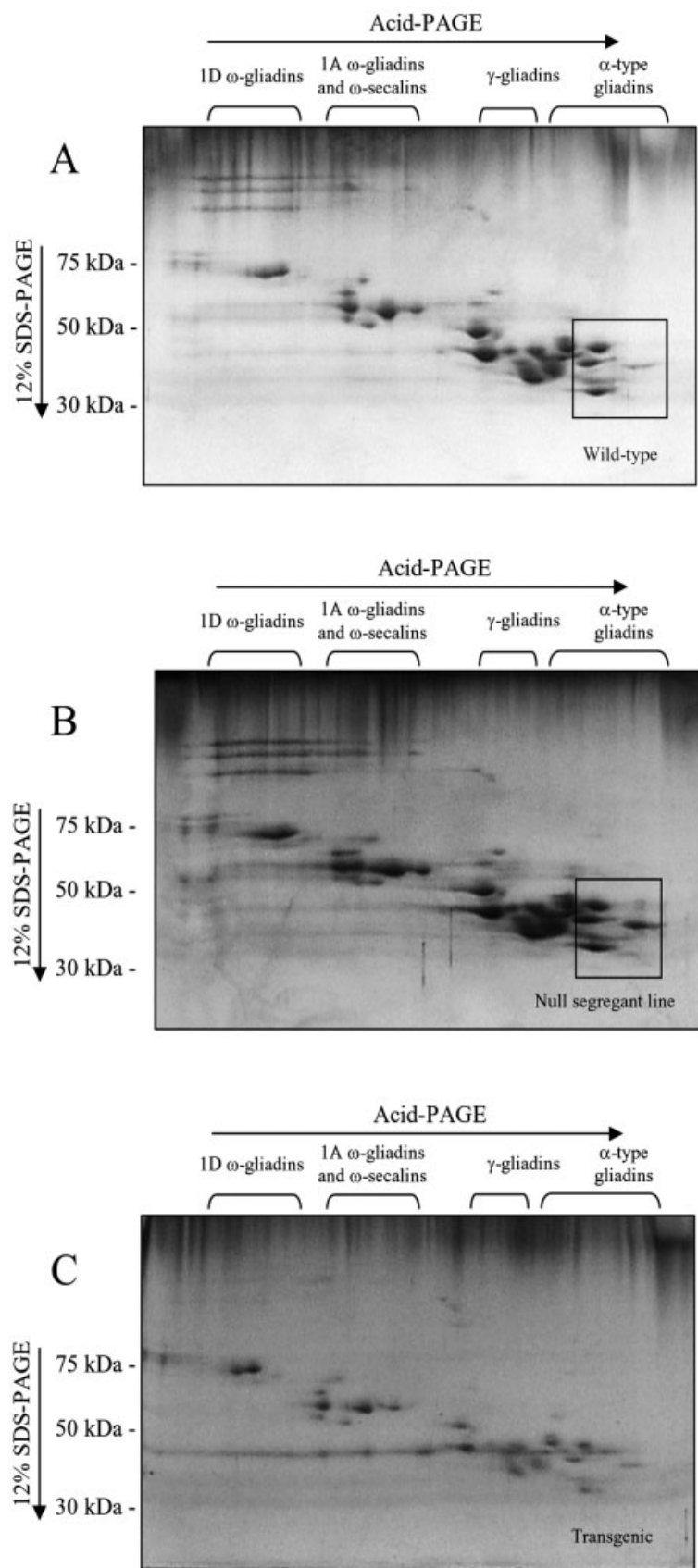


**Figure 5.** Deduced amino acid sequence of the transgenic LMW-GS encoded by the clone F23A [26]. Underlined or overlined are fragments identified by N-terminal sequencing in C-type LMW-GS preparations from the transgenic genotype. Numbers shown on the sequence correspond to the spots as numbered in Fig. 4C; their sequences are reported in Table 4. Vertical brackets and directed arrows define the borders and the start positions of the N-terminal, central repetitive and C-terminal domains of a typical LMW-GS, respectively.

### 3.2.1.3 Gliadins

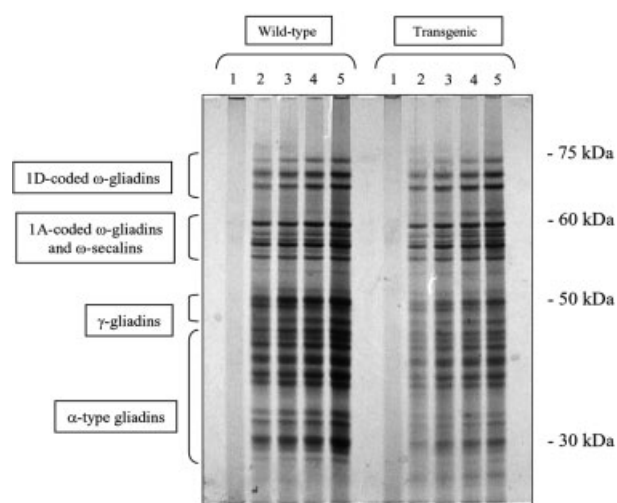
Based on our microarray data, all classes of gliadin genes were heavily down-regulated during seed development in the transgenic line (Figs. 3C–E). With the aim of verifying the concordance between transcript and protein abundances, gliadins were investigated by means of selective extraction and 2-DE. We have used Acid-PAGE vs. SDS-PAGE, instead of IEF vs. SDS-PAGE, in order to easily recognize the different gliadin classes that are typically classified according to their mobility in Acid-PAGE [3]. In Fig. 6, the 2-DE maps of total gliadin preparations from the wild-type, null segregant line and transgenic genotype extracted from mature grains are shown. The 2-D gels were loaded with the same volume of protein extracts derived from equal masses of seeds. In general, gliadins were much less abundant in the transgenic than in the wild-type genotype. After expression analysis of 2-D gels, a group of  $\alpha$ -type gliadins was regarded as significantly differentially synthesized in the wild-type with respect to the transgenic genotype (boxed in Fig. 6A). The identity of these spots was confirmed by N-terminal sequencing of the first eight amino acids (data not shown). The null segregant line was not significantly different from the wild-type cultivar (Fig. 6B), thus the down-regulation of gliadins, and, in particular of  $\alpha$ -type gliadins, was exclusive to the transgenic genotype.

According to the transcript profiles (Figs. 3C–E), the differential expression of gliadin genes was established early in seed development. Several gliadin-encoding clones were already significantly down-regulated at 10 DPA late in the transgenic line. We thus investigated the patterns of accu-



**Figure 6.** 2-DE (Acid-PAGE × SDS-PAGE) of a gliadin preparation from mature grains of the wild-type (A), null segregant line (B), and transgenic genotype (C). Equal volumes of seed extracts derived from equal masses of seeds were loaded on the gels. Boxed spots were significantly expressed by the wild-type and null segregant line and were submitted to N-terminal sequencing.

mulation of gliadins during the period of seed maturation. The 1-D electrophoretic patterns of gliadin extracts from the wild-type and transgenic genotypes during grain development is illustrated in Fig. 7. Also in this case, the 1-D gels were loaded with the same volume of protein extracts derived from equal masses of seeds. It is interesting to note that  $\alpha$ -type and  $\gamma$ -gliadins electrophoretic patterns are less intense in the transgenic compared to the control line. This is valid also for  $\omega$ -gliadins (as well as  $\omega$ -secalins), although to a lower extent compared to the other gliadin classes. These results parallel the microarray data, therefore transcripts and proteins accumulation are concurrent, and the observed difference between the various gliadin classes might mean that they are under different controls, but this has not been investigated in this context. In general, lower levels of gliadins have also been reported recently in transgenic wheat overexpressing the HMW-GS 1Dx5 [65].



**Figure 7.** 1-D Acid-PAGE separation of total gliadin extracts from the developing seeds of the wild-type (left) and transgenic genotype (right). Equal volumes of seed extracts derived from equal masses of seeds were loaded on the gels. 1, 10 DPA late; 2, 20 DPA; 3, 30 DPA; 4, 35 DPA; 5, mature grain (around 45 DPA). The presence of  $\omega$ -secalins is due to the 1B-1R translocation carried by the bread wheat cultivar Bobwhite.

### 3.2.2 Metabolic proteins

In addition to the gluten proteins, we also analyzed the KCl-soluble fraction designated as “metabolic proteins” [14]. Besides gluten proteins, which are toxic for certain individuals (*i.e.*, celiac disease), wheat grains contain a high concentration of potentially allergenic polypeptides [66] that mostly belong to this fraction.

The metabolic fraction resolved into multiple spots in the 2-D gel as shown in Fig. 8. In terms of total protein spot volumes, the transgenic genotype has a mean reduction in global intensity of 32% with respect to the wild-type genotype ( $t$ -test  $p$ -value < 0.01).

**Table 5.** DE metabolic proteins

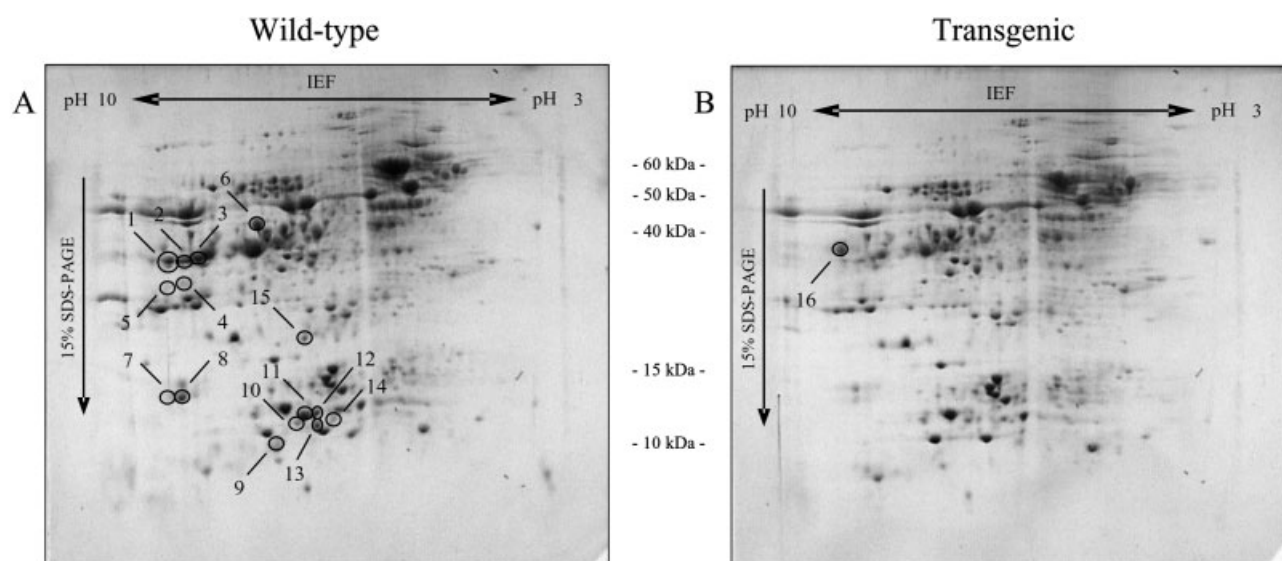
Spot no.	NCBI accession number	Protein name
1	AAM88383	Peroxidase 1
2	AAM88383	Peroxidase 1
3	AAM88383	Peroxidase 1
4	ABG68030	Globulin 1
5	AAP80642	19 kDa globulin
6	AAK84773	$\gamma$ -Gliadin
7	AAP06762	Grain softness protein-1
8	AAK84773	$\gamma$ -Gliadin
9	AAP80612	Unknown
10	P16850	Wheat $\alpha$ -amylase trypsin inhibitor CM1
11	AAY42618	0.19 dimeric $\alpha$ -amylase inhibitor
12	AAR28678	$\alpha$ -Amylase inhibitor CM3
13	AAV39517	0.19 dimeric $\alpha$ -amylase inhibitor
14	CAA39099	CM2 protein
15	BAC76688	27K protein
16	AAO53265	LMW-GS

MS/MS identifications of the DE metabolic proteins at the level of mature grains between the transgenic and the wild-type genotype. The amino acid sequences of the peptides identified for each protein and the percentage of the protein coverage that they represent are provided along with other details in Supporting Information Table 5. Identified proteins are based on the criteria specified for protein identification in Section 2.

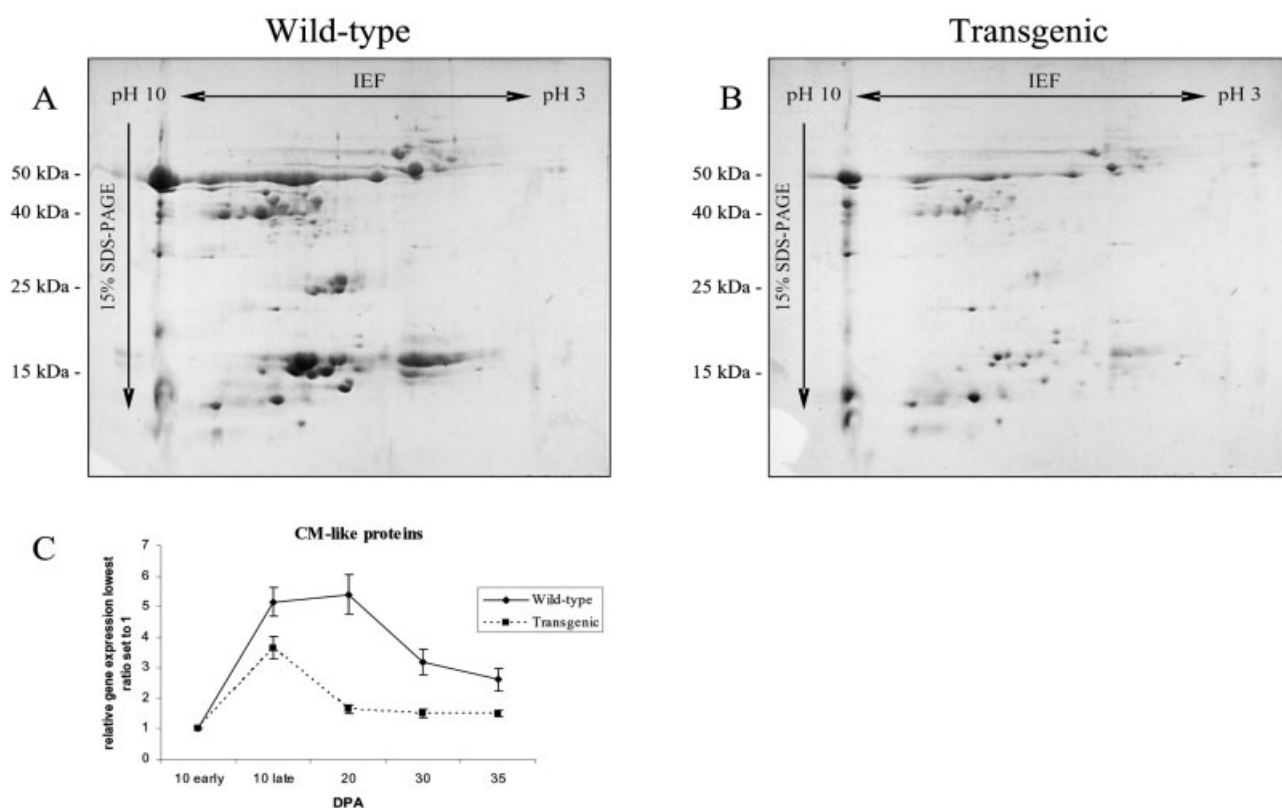
Expression analysis of 2-D gels, performed to identify individual differentially accumulated metabolic proteins, identified several spots that were selected for MS/MS analyses. Table 5 reports the corresponding protein identifications. All the proteins identified were significantly down-regulated in the transgenic genotype, except for up-regulated spot 16 that corresponds to the contaminant transgenic LMW-GS.

Spots 1, 2, and 3 (Table 5) were identified as seed peroxidases. Evidence of differential synthesis of peroxidases also came from microarray analyses, where a clone encoding a seed peroxidase was found to be constantly down-regulated in the transgenic genotype (clone BE590763, see Supporting Information Appendix S3 Transcript Profiles).

Spots 10–14 were all CM-like proteins or inhibitors of  $\alpha$ -amylase/trypsin. The various protease and amylase inhibitors of wheat seeds all belong to the family of CM-like proteins, and are proteins usually synthesized in the cereal seeds to provide a defense against the protease enzymes secreted by pathogens [67, 68]. Attributions of the remaining spots are reported in Table 5. Since some of the spots identified as differentially synthesized in metabolic protein preparations were actually CM-like proteins, we decided to investigate their abundance in mature seeds by means of selective extraction followed by analysis of their 2-DE gel pattern (Figs. 9A and B). The transgenic line shows a drastic decrease in the amount of all the CM-like proteins (78%,  $t$ -test



**Figure 8.** Analysis of the metabolic protein fraction: 2-DE (IEF  $\times$  SDS-PAGE, pH 3–10) of metabolic proteins from mature endosperms of wild-type (A) and transgenic seeds (B). Equal volumes of seed extracts derived from equal masses of seeds were loaded on the gels. The numbered spots (DE spots) were characterized by MS/MS analyses (Table 5).



**Figure 9.** Analysis of the CM-like proteins: 2-DE (IEF  $\times$  SDS-PAGE, pH 3–10) of the CM-like proteins from mature endosperms of the wild-type (A) and transgenic seeds (B). Equal volumes of seed extracts derived from equal masses of seeds were loaded on the gels. (C) reports the “pooled” expression profiles of several clones encoding CM-like proteins. Expression is related to 10 DPA early and the ratios were rescaled by setting the lowest ratio to 1. At each time-point, the ratio between the relative gene expression values is the fold change between the wild-type and the transgenic genotype. In (C), vertical bars denote the SEM. Where vertical bars are not visible, they are smaller than the data points.

$p$ -value<0.01), with respect to the wild-type genotype. This decrease, which was also detectable at the transcript level in developing seeds for several clones encoding CM-like proteins (Fig. 9C), can again be interpreted as an extreme compensatory effect to the high overexpression of the transgenic LMW glutenin subunit gene. CM-like proteins are in fact distantly related to the wheat prolamins [69], and show a high degree of sequence identity with the nonrepetitive domains typically present in the prolamins sequences. As with the HMW-GS and gliadins, the decrease in the abundance of CM-like proteins could be likely a consequence of the diversion of the amino acids pool to the synthesis of the transgenic glutenin.

#### 4 Concluding remarks

The experiments described here provide a global evaluation at the transcriptome and proteome level of the consequence of a biolistic transformation event in bread wheat that leads to an unusually high expression of an LMW-GS transgene. Wheat cDNA microarrays and 2-DE demonstrated their high and complementary potential to detect significant changes, and allowed us to assess the concordance between the RNA and protein levels in the developing seeds.

Referring to the definitions provided in ref. [70], the transgenic line analyzed here showed a series of “unintended effects.” As a response to the overaccumulation of the LMW glutenin, the transgenic seeds have adopted a global compensatory response consisting in a heavy down-regulation of all other classes of storage proteins. This response might be directed to maintain a homeostatic level of seed storage proteins compatible with seed development and germination. It is known that plants might respond to a drastic increase in the production of nutrients or storage compounds, by activating a silencing/degradation pathway that negates the overaccumulation of storage reserves [71, 72].

Since gene expression changes during seed development, although in different directions, we assume that the genes encoding gliadins, glutenins, and CM-like proteins (related to prolamins) [73] are all likely subjected to a similar transcriptional regulation. In this line, the compensatory responses started immediately following the early accumulation of the transgenic glutenin and persisted until seed maturation. The observed compensation could likely be a consequence of the diversion of the amino acid reserves to the synthesis of the transgenic glutenin. The magnitude of these compensatory effects seems to be proportional to the expression level of the introduced glutenin transgenes [60, 61, 65].

Similarly, also the timing of the accumulation of the transgenic glutenin during seed development constitutes an unintended effect. It is in fact possible that this transgene, encoded at the *Glu-D3* locus, belongs to a class “early expressors” LMW-GS genes. Early expressors  $\alpha$ -gliadin genes from the D genome have been identified by a large-scale *in*

*silico* analysis of wheat ESTs [74]. The possible lack of specific regulatory sequences, upstream and/or downstream the coding sequence, may also exert an influence on the timing of accumulation of the transgenic transcript. Moreover, the simultaneous regulation of the transgenic LMW-GS, both at the transcriptional and protein synthesis level, may be explained on the basis of the cointegration of the transgenes in a single site [24].

Interestingly, the development of novel allergenic compounds or toxic substances or the increase of those already present as a result of genetic modification is a central issue in the assessment of any GM crops [75]. It is worth mentioning that, among the “predictable unintended effects”, the transgenic seeds also exhibited globally a lesser amount of polypeptides with allergenic potential (gliadins and CM-like proteins) [76, 77].

This paper follows several reports investigating the transcriptional, proteomic or metabolomic differences between GM plants and their corresponding traditional varieties [78–82]. Because the work here reported is based on the analysis of a single line, it is not possible to draw general conclusions about the substantial equivalence of this GM genotype, also because a significant proportion of the DE clones corresponded to unknown genes. However, most of the variations observed include predictable alterations of seed transcriptome and proteome. These were for the great part compensatory effects limited mostly to the prolamins superfamily, which includes seed storage proteins along with defense related proteins.

The data here obtained offer the possibility to speculate about the organization of the gluten matrix, whose detailed structure is still poorly understood [4]. The presence of polypeptide fragments corresponding to partial sequences of the transgene might be the consequence of the insertion of transgene fragments, which are eventually translated, but could also originate from the proteolytic cleavage of the transgenic glutenin. In our previous work [23], it was shown that the majority of glutenin polymers in the transgenic line were actually homopolymers of the transgenic protein itself. These “unusual” polymers, whose accumulation may also constitute a stress for the ER [83], may be recognized as incorrect by the cell, thus triggering the degradation process.

In conclusion, although this work is based on the analysis of a single transgenic line, this genotype gives important clues that help in understanding mechanisms controlling the synthesis, the organization of wheat seed proteins, and the expression of their coding genes during seed development.

*This work was supported by the Italian Minister for University and Research (MIUR), projects “Gene expression and accumulation of agronomically important proteins in plant cells: transcriptional and post-transcriptional mechanisms” (FIRB RBNE01TYZF), and “Evaluation of the presence of potentially nutritional and/or allergenic proteins in commercial and trans-*

genic wheat endosperms and their characterization" (PRIN 2006). The authors wish to acknowledge Dr. D. D. Kasarda (USDA, Albany, CA) for his stimulating discussions, Dr. A. E. Blechl for the critical revision of the manuscript, D. Ferri and E. Cannarella (University of Tuscia, Viterbo, Italy) for technical assistance. Mention of products by USDA-ARS does not imply endorsement to the exclusion of others that may also be suitable.

The authors have declared no conflict of interest.

## 5 References

- [1] Shewry, P. R., Jones, H. D., Transgenic wheat: Where do we stand after the first 12 years? *Ann. Appl. Biol.* 2005, **147**, 1–14
- [2] Shewry, P. R., Halford, N. G., Lafiandra, D., Genetics of wheat gluten proteins. *Adv. Genet.* 2003, **49**, 111–184
- [3] Woychik, J. H., Boundy, J. A., Dimler, R. J., Starch gel electrophoresis of wheat gluten proteins with concentrated urea. *Arch. Biochem. Biophys.* 1961, **94**, 477–482
- [4] Wrigley, C. W., Giant proteins with flour power. *Nature* 1996, **381**, 738–739
- [5] Jackson, E. A., Holt, L. M., Payne, P. I., Characterisation of high-molecular-weight gliadin and low-molecular-weight glutenin subunits of wheat endosperm by two-dimensional electrophoresis and chromosomal localisation of their controlling genes. *Theor. Appl. Genet.* 1983, **66**, 29–37
- [6] D'Ovidio, R., Masci, S., The low-molecular-weight glutenin subunits of wheat gluten. *J. Cereal Sci.* 2004, **39**, 321–339
- [7] Masci, S., Rovelli, L., Kasarda, D. D., Vensel, W. H., Lafiandra, D., Characterisation and chromosomal localisation of C-type low-molecular-weight glutenin subunits in the bread wheat cultivar Chinese Spring. *Theor. Appl. Genet.* 2002, **104**, 422–428
- [8] Masci, S., Lafiandra, D., Porceddu, E., Lew, E. J. L. *et al.*, D-glutenin subunits: N-terminal sequences and evidence for the presence of cysteine. *Cereal Chem.* 1993, **70**, 581–585
- [9] Masci, S., Egorov, T. A., Ronchi, C., Kuzmicky, D. D. *et al.*, Evidence for the presence of only one cysteine residue in the D-type low molecular weight subunits of wheat glutenin. *J. Cereal Sci.* 1999, **29**, 17–25
- [10] Altenbach, S. B., Kothari, K. M., Omega gliadin genes expressed in *Triticum aestivum* cv Butte 86: Effects of post-anthesis fertilizer on transcript accumulation during grain development. *J. Cereal Sci.* 2007, **46**, 169–177
- [11] Shewry, P. R., Lafiandra, D., Salcedo, G., Aragoncillo, C. *et al.*, N-terminal amino acid sequences of chloroform/methanol-soluble proteins and albumins from endosperms of wheat, barley and related species. *FEBS Lett.* 1984, **175**, 359–363
- [12] Barber, D., Sanchez-Monge, R., Garcia-Olmedo, F., Salcedo, G., Mendez, E., Evolutionary implications of the homologies among members of the trypsin/ $\alpha$ -amylase inhibitor family (CM-proteins) in wheat and barley. *Biochim. Biophys. Acta* 1986, **873**, 145–151
- [13] Wong, J. H., Cai, N., Tanaka, C. K., Vensel, W. H. *et al.*, Thio-redoxin reduction alters the solubility of proteins of wheat starchy endosperm: An early event in cereal germination. *Plant. Cell. Physiol.* 2004, **45**, 407–415
- [14] Hurkman, W. J., Tanaka, C. K., Improved methods for separation of wheat endosperm proteins and analysis by two-dimensional gel electrophoresis. *J. Cereal Sci.* 2004, **40**, 295–299.
- [15] Hurkman, W. J., Tanaka, C. K., Extraction of wheat endosperm proteins for proteome analysis. *J. Chromatogr. B* 2007, **849**, 344–350
- [16] Vensel, W. H., Tanaka, C. K., Cai, N., Wong, J. H. *et al.*, Developmental changes in the metabolic protein profiles of wheat endosperm. *Proteomics* 2005, **5**, 1594–1611
- [17] Altpeter, F., Vasil, V., Srivastava, V., Vasil, I. K., Integration and expression of the high-molecular-weight glutenin subunit 1Ax1 gene into wheat. *Nat. Biotechnol.* 1996, **14**, 1151–1159
- [18] Blechl, A. E., Anderson, O. D., Expression of a novel high-molecular-weight glutenin subunit gene in transgenic wheat. *Nat. Biotechnol.* 1996, **14**, 875–879
- [19] Tosi, P., D'Ovidio, R., Napier, J. A., Bekes, F., Shewry, P. R., Expression of epitope-tagged LMW glutenin subunits in the starchy endosperm of transgenic wheat and their incorporation into glutenin polymers. *Theor. Appl. Genet.* 2004, **108**, 468–476
- [20] Tosi, P., Masci, S., Giovangrossi, A., D'Ovidio, R. *et al.*, Modification of the low molecular weight (LMW) glutenin composition of transgenic durum wheat: Effects on glutenin polymer size and gluten functionality. *Mol. Breed.* 2005, **16**, 113–126
- [21] Darlington, H., Fido, R., Tatham, A. S., Jones, H. *et al.*, Milling and baking properties of field grown wheat expressing HMW subunit transgenes. *J. Cereal Sci.* 2003, **38**, 301–306
- [22] Gupta, R. B., Singh, N. K., Shepherd, K. W., The cumulative effect of allelic variation in LMW and HMW glutenin subunits on physical dough properties in progeny of two bread wheats. *Theor. Appl. Genet.* 1989, **77**, 57–64
- [23] Masci, S., D'Ovidio, R., Scossa, F., Patacchini, C. *et al.*, Production and characterization of a transgenic bread wheat line over-expressing a low-molecular-weight glutenin subunit gene. *Mol. Breed.* 2003, **12**, 209–222
- [24] Carozza, R., Scossa, F., Masci, S., D'Ovidio, R., Ceoloni, C., 2005 FISH-based physical mapping of a low-molecular-weight glutenin subunit transgene insertion in a transformed bread wheat genotype over-expressing the transgene-encoded proteins. *Proc. XLIX Ann. S.I.G.A. Congr.*, [http://www.siga.unina.it/SIGA\\_2005/SIGA2005/B\\_28.pdf](http://www.siga.unina.it/SIGA_2005/SIGA2005/B_28.pdf).
- [25] Altenbach, S. B., DuPont, F. M., Kothari, K. M., Chan, R. *et al.*, Temperature, water and fertilizer influence the timing of key events during grain development in a US spring wheat. *J. Cereal Sci.* 2003, **37**, 9–20
- [26] Cassidy, B. G., Dvorak, J., Anderson, O. D., The wheat low-molecular-weight glutenin genes: Characterization of six new genes and progress in understanding gene family structure. *Theor. Appl. Genet.* 1998, **96**, 743–750
- [27] Laudencia-Chingcuanco, D. L., Stamova, B., Lazo, G. R., Cui, X., Anderson, O. D., Analysis of the wheat endosperm transcriptome. *J. Appl. Genet.* 2006, **47**, 287–302
- [28] Zhang, D., Choi, D. W., Wanamaker, S., Fenton, R. D. *et al.*, Construction and evaluation of cDNA libraries for large-scale expressed sequence tag sequencing in wheat (*Triticum aestivum* L.). *Genetics* 2004, **168**, 595–608
- [29] Qi, L. L., Echaliier, B., Chao, S., Lazo, G. R. *et al.*, A chromosome bin map of 16,000 expressed sequence tag loci and

- distribution of genes among the three genomes of polyploid wheat. *Genetics* 2004, **168**, 701–712
- [30] Kerr, M. K., Design considerations for efficient and effective microarray studies. *Biometrics* 2003, **59**, 822–828
- [31] Ruuska, S. A., Girke, T., Benning, C., Ohlrogge, J. B., Contrapuntal networks of gene expression during Arabidopsis seed filling. *Plant. Cell* 2002, **14**, 1191–1206
- [32] Laudencia-Chingcuanco, D. L., Stamova, B. S., You, F. M., Lazo, G. R. *et al.*, Transcriptional profiling of wheat caryopsis development using cDNA microarrays. *Plant Mol. Biol.* 2006, **63**, 651–668
- [33] Smyth, G. K., Speed, T. P., Normalization of cDNA microarray data. *Methods* 2003, **31**, 265–273
- [34] Smyth, G. K., Limma: Linear models for microarray data. in: Gentleman, R., Carey, V., Dudoit, S., Irizarry, R., Huber, W. (Eds.), *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, Springer, New York 2005, pp. 397–420
- [35] Ihaka, R., Gentleman, R., R: A language for data analysis and graphics. *J. Comput. Graph. Statist.* 1996, **5**, 299–314
- [36] Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B. *et al.*, Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biol.* 2004, **5**, R80
- [37] Lönstedt, I., Speed, T., Replicated microarray data. *Stat. Sin.* 2002, **12**, 31–46
- [38] Benjamini, Y., Hochberg, V., Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 1995, **57**, 289–300
- [39] Galbraith, D. W., Global analysis of cell type-specific gene expression. *Comp. Funct. Genomics* 2003, **4**, 208–215
- [40] Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G. *et al.*, Minimum information about a microarray experiment (MIAME) - toward standards for microarray data. *Nat. Genet.* 2001, **29**, 365–371.
- [41] Østergaard, O., Finnie, C., Laugesen, S., Roepstorff, P., Svensson, B., Proteome analysis of barley seeds: Identification of major proteins from two-dimensional gels (pI 4–7). *Proteomics* 2004, **4**, 2437–2447
- [42] Neuhoff, V., Arold, N., Taube, D., Ehrhardt, W., Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with a clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* 1988, **9**, 255–262
- [43] Khan, K., Hamada, A. S., Patek, J., Polyacrylamide gel electrophoresis for wheat variety identification: Effect of variables on gel properties. *Cereal Chem.* 1985, **62**, 310–313
- [44] Craig, R., Beavis, R. C., TANDEM: Matching proteins with tandem mass spectra. *Bioinformatics* 2004, **20**, 1466–1467
- [45] Keller, A., Nesvizhskii, A. I., Kolker, E., Aebersold, R., Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* 2002, **74**, 5383–5392
- [46] Nesvizhskii, A. I., Keller, A., Kolker, E., Aebersold, R. A., statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* 2003, **75**, 4646–4658
- [47] Halford, N. G., Shewry, P. R., The structure and expression of cereal storage protein genes. *Plant Cell Monographs* 2007, **8**, 195–218
- [48] Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, A. A. *et al.*, Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucl. Acids Res.* 1997, **25**, 3389–3402
- [49] Bowles, D., Defense-related proteins in higher plants. *Ann. Rev. Biochem.* 1990, **59**, 873–907
- [50] Wasternack, C., Parthier, B., Jasmonate-signalled plant gene expression. *Trends Plant Sci.* 1997, **2**, 302–307
- [51] Altenbach, S. B., Kothari, K. M., Lieu, D., Environmental conditions during wheat grain development alter temporal regulation of major gluten protein genes. *Cereal Chem.* 2002, **79**, 279–285
- [52] Gupta, R. B., Masci, S., Lafiandra, D., Bariana, H. S., MacRitchie, F., Accumulation of protein subunits and their polymers in developing grains of hexaploid wheats. *J. Exp. Bot.* 1996, **47**, 1377–1385
- [53] Jones, H. D., Wu, H., Sparks, C., Doherty, A. *et al.*, The potential of biotechnology to produce novel gluten phenotypes. in: Lafiandra, D., Masci, S., D'Ovidio, R. (Eds.), *The Gluten Proteins*, Royal Society of Chemistry, Cambridge, pp. 14–17
- [54] Carceller, J. L., Aussenac, T., SDS-insoluble glutenin polymer formation in developing grains of hexaploid wheat: The role of the ratio of high to low molecular weight glutenin subunits and drying rate during ripening. *Aust. J. Agric. Res.* 2001, **28**, 193–201
- [55] DuPont, F. M., Chan, R., Lopez, R., Vensel, W. H., Sequential extraction and quantitative recovery of gliadins, glutenins, and other proteins from small samples of wheat flour. *J. Agric. Food Chem.* 2005, **53**, 1575–1584
- [56] Barro, F., Rooke, L., Bekes, F., Gras, P. *et al.*, Transformation of wheat with high molecular weight subunit genes results in improved functional properties. *Nat. Biotechnol.* 1997, **15**, 1295–1299
- [57] Rooke, L., Bekes, F., Fido, R., Barro, F. *et al.*, Overexpression of a gluten protein in transgenic wheat results in greatly increased dough strength. *J. Cereal Sci.* 1999, **30**, 115–120
- [58] Alvarez, M. L., Guelman, S., Halford, N. G., Lustig, S. *et al.*, Silencing of HMW glutenins in transgenic wheat expressing extra HMW subunits. *Theor. Appl. Genet.* 2000, **100**, 319–327
- [59] Vasil, I. K., Bean, S., Zhao, J., McCluskey, P. *et al.*, Evaluation of baking properties and gluten protein composition of field grown transgenic wheat lines expressing high molecular weight glutenin gene 1Ax1. *J. Plant Physiol.* 2001, **158**, 521–528
- [60] He, G. Y., Jones, H. D., D'Ovidio, R., Masci, S. *et al.*, Expression of an extended HMW subunit in transgenic wheat and the effect on dough mixing properties. *J. Cereal Sci.* 2005, **42**, 225–231
- [61] Blechl, A. E., Lin, J., Nguyen, S., Chan, R. *et al.*, Transgenic wheats with elevated levels of Dx5 and/or Dy10 high-molecular-weight glutenin subunits yield doughs with increased mixing strength and tolerance. *J. Cereal Sci.* 2007, **45**, 172–183
- [62] Doran, P. M., Foreign protein degradation and instability in plants and plant tissue cultures. *Trends Biotechnol.* 2006, **24**, 426–432
- [63] Dolja, V. V., Peremyslov, V. V., Keller, K. E., Martin, R. R., Hong, J., Isolation and stability of histidine-tagged proteins produced in plants via potyvirus gene vectors. *Virology* 1998, **252**, 269–274

- [64] Smith, M. L., Keegan, M. E., Mason, H. S., Shuler, M. L., Factors important in the extraction stability and in vitro assembly of the hepatitis B surface antigen derived from recombinant plant systems. *Biotechnol. Prog.* 2002, 18, 538–550
- [65] Rakszegi, M., Békés, F., Láng, L., Tamás, L. *et al.*, Technological quality of transgenic wheat expressing an increased amount of a HMW glutenin subunit. *J. Cereal Sci.* 2005, 42, 15–23
- [66] Breiteneder, H., Mills, E. N. C., Molecular properties of food allergens. *J. Allergy Clin. Immunol.* 2005, 115, 14–23
- [67] Gutierrez, C., Sanchez-Monge, R., Gomez, L., Ruiz-Tapiador, M. *et al.*,  $\alpha$ -Amylase activities of agricultural insect pests are specifically affected by different inhibitor preparations from wheat and barley endosperms. *Plant Sci.* 1990, 72, 37–44
- [68] Finnie, C., Melchior, S., Roepstorff, P., Svensson, B., Proteome analysis of grain filling and seed maturation in barley. *Plant Physiol.* 2002, 129, 1308–1319
- [69] Anderson, O. D., Hsia, C. C., Adalsteins, A. E., Lew, E. J. L., Kasarda, D. D., Identification of several new classes of low-molecular-weight wheat gliadin-related proteins and genes. *Theor. Appl. Genet.* 2001, 103, 307–315
- [70] Cellini, F., Chesson, A., Colquhoun, I., Constable, A. *et al.*, Unintended effects and their detection in genetically modified crops. *Food Chem. Toxicol.* 2004, 42, 1089–1125
- [71] Karchi, H., Shaul, O., Galili, G., Lysine synthesis and catabolism are coordinately regulated during tobacco seed development. *Proc. Natl. Acad. Sci. USA* 1994, 91, 2577–2581
- [72] Tang, G., Galili, G., Using RNAi to improve plant nutritional value: From mechanism to application. *Trends Biotechnol.* 2004, 22, 463–469
- [73] Shewry, P. R., Plant storage proteins. *Biol. Rev.* 1995, 70, 375–426
- [74] Kawaura, K., Mochida, K., Ogiwara, Y., Expression profile of two storage-protein gene families in hexaploid wheat revealed by large-scale analysis of expressed sequence tags. *Plant Phys.* 2005, 139, 1870–1880
- [75] FAO/WHO *Allergenicity of genetically modified foods. Report of a Joint FAO/WHO Expert Consultation on Foods Derived from Biotechnology*, Rome, Italy 2001.
- [76] Mills, E. N. C., Madsen, C., Shewry, P. R., Wichers, H. J., Food allergens of plant origin – their molecular and evolutionary relationships. *Trends Food Sci. Technol.* 2003, 14, 145–156
- [77] Breiteneder, H., Radauer, C., A classification of plant food allergens. *J. Allergy Clin. Immunol.* 2004, 113, 821–830
- [78] Gregersen, P. L., Brinch-Pedersen, H., Holm, P. B., A microarray-based comparative analysis of gene expression profiles during grain development in transgenic and wild type wheat. *Transgenic Res.* 2005, 14, 887–905
- [79] Lehesranta, S. J., Davies, H. V., Shepherd, V. T., Nunan, N. *et al.*, Comparison of tuber proteomes of potato varieties landraces and genetically modified lines. *Plant Physiol.* 2005, 138, 1690–1699
- [80] Baker, J. M., Hawkins, N. D., Ward, J. L., Lovegrove, A. *et al.*, A metabolomic study of substantial equivalence of field-grown genetically modified wheat. *Plant Biotechnol. J.* 2006, 4, 381–392
- [81] Baudo, M. M., Lyons, R., Powers, S., Pastori, G. M. *et al.*, Transgenesis has less impact on the transcriptome of wheat grain than conventional breeding. *Plant Biotechnol. J.* 2006, 4, 369–380
- [82] Ruebelt, M. C., Lipp, M., Reynolds, T. L., Schmukey, J. J. *et al.*, Application of two-dimensional gel electrophoresis to interrogate alterations in the proteome of genetically modified crops. 3. Assessing unintended effects. *J. Agric. Food Chem.* 2006, 54, 2169–2177
- [83] Urade, R., Cellular response to unfolded proteins in the endoplasmic reticulum of plants. *FEBS J.* 2007, 274, 1152–1171